

=====

Sequence Listing was accepted.

If you need help call the Patent Electronic Business Center at (866)  
217-9197 (toll free).

Reviewer: Durreshwar Anjum

Timestamp: Wed Oct 24 12:54:29 EDT 2007

=====

Application No: 10622220 Version No: 2.0

Input Set:

Output Set:

Started: 2007-10-05 10:38:12.131  
Finished: 2007-10-05 10:38:13.684  
Elapsed: 0 hr(s) 0 min(s) 1 sec(s) 553 ms  
Total Warnings: 8  
Total Errors: 0  
No. of SeqIDs Defined: 20  
Actual SeqID Count: 20

| Error code | Error Description                                   |
|------------|---|
| W 213      | Artificial or Unknown found in <213> in SEQ ID (13) |
| W 213      | Artificial or Unknown found in <213> in SEQ ID (14) |
| W 213      | Artificial or Unknown found in <213> in SEQ ID (15) |
| W 213      | Artificial or Unknown found in <213> in SEQ ID (16) |
| W 402      | Undefined organism found in <213> in SEQ ID (17)    |
| W 402      | Undefined organism found in <213> in SEQ ID (18)    |
| W 213      | Artificial or Unknown found in <213> in SEQ ID (19) |
| W 213      | Artificial or Unknown found in <213> in SEQ ID (20) |

# SEQUENCE LISTING

<110> Nilles, Matthew L.  
Matson, Jyl S.

<120> YERSINIA SPECIES COMPOSITIONS

<130> 3128-6046US

<140> 10622220

<141> 2003-07-18

<150> US 60/444,076

<151> 2003-01-31

<160> 20

<170> PatentIn version 3.2

<210> 1

<211> 87

<212> PRT

<213> Yersinia pestis

<220>

<223> amino acid sequence of YscF

<220>

<223> sequence can be found at MedLine accession number NP\_857921.1

<400> 1

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu  
1 5 10 15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys  
20 25 30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro  
35 40 45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile  
50 55 60

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln  
65 70 75 80

Gly Ile Leu Gln Lys Phe Pro  
85

<210> 2

<211> 264

<212> DNA  
 <213> *Yersinia pestis*

<220>  
 <223> nucliec acids encoding YscF

<220>  
 <223> sequence can be found at MedLine accession number NL\_004839

<400> 2

|   |     |
|---|-----|
| atgagtaact tctctggatt tacgaaagga accgatatcg cagacttaga tgcggtggct | 60  |
| caaacgctca agaagccagc agacgatgca aacaaagcgg ttaatgactc gatagcagca | 120 |
| ttgaaagata agcctgacaa cccggcgcta cttgctgact tacaacattc aattaataaa | 180 |
| tggtcggttaa ttacaatat aaactcaacc atagttcgta gcatgaaaga cttaatgcaa | 240 |
| ggcatcctac agaagttccc ataa  | 264 |

<210> 3  
 <211> 264  
 <212> DNA  
 <213> *Yersinia enterocolitica*

<220>  
 <223> nucleic acid encoding YscF from Y. enterocolitica

<220>  
 <223> sequence can be found at MedLine accession number NC\_005017

<400> 3

|   |     |
|---|-----|
| atgagtaatt tctctgggtt tgcaaaagga accgatatca cagacttaga tgcggtggct | 60  |
| caaacgctca agaagccagc agacggcgca aacaaagcgg ttaatgactc gatagcagcg | 120 |
| ttgaaagaga cgctgacaa cccggcgcta cttgctgact tacaacattc aattaataaa  | 180 |
| tggtcggttaa ttacaatat aagctcaacc atagttcgta gcatgaaaga cttaatgcaa | 240 |
| ggcatcctac agaagttccc ataa  | 264 |

<210> 4  
 <211> 87  
 <212> PRT  
 <213> *Yersinia enterocolitica*

<220>  
 <223> amino acid sequence of YscF from Y. enterocolitica

<220>  
 <223> sequence can be found at MedLine accession number NP\_863538.1

<400> 4

Met Ser Asn Phe Ser Gly Phe Ala Lys Gly Thr Asp Ile Thr Asp Leu  
1 5 10 15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys  
20 25 30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Thr Pro Asp Asn Pro  
35 40 45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile  
50 55 60

Tyr Asn Met Ser Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln  
65 70 75 80

Gly Ile Leu Gln Lys Phe Pro  
85

<210> 5

<211> 80

<212> PRT

<213> Salmonella enterica

<220>

<223> amino acid sequence of PrgI

<220>

<223> sequence can be found at MedLine accession number CAD05980.1

<400> 5

Met Pro Thr Ser Trp Ser Gly Tyr Leu Asp Glu Val Ser Ala Lys Phe  
1 5 10 15

Asp Lys Gly Val Asp Asn Leu Gln Thr Gln Val Thr Glu Ala Leu Asp  
20 25 30

Lys Leu Ala Ala Lys Pro Ser Asp Pro Ala Leu Leu Ala Ala Tyr Gln  
35 40 45

Ser Lys Leu Ser Glu Tyr Asn Leu Tyr Arg Asn Ala Gln Ser Asn Thr  
50 55 60

Val Lys Val Phe Lys Asp Ile Asp Ala Ala Ile Ile Gln Asn Phe Arg  
65 70 75 80

<210> 6  
<211> 243  
<212> DNA  
<213> Salmonella enterica

<220>  
<223> sequence can be found at MedLine accession number AL627276

<220>  
<223> nucleic acid encoding PrgI

<400> 6  
atgccaacat cttggtcagg ctatctggat gaagtttcag caaaatttga taagggcggt 60  
gataatctac aaacgcaggt aacagaggcg ctggataaat tagcagcaaa accctccgat 120  
ccggcgctac tggcggcgta tcagagtaag ctctcggaat ataacttgta ccgtaacgcg 180  
caatcgaaca cggtaaaagt ctttaaggat attgatgctg ccattattca gaacttccgt 240  
taa 243

<210> 7  
<211> 83  
<212> PRT  
<213> Shigella flexneri

<220>  
<223> amino acid sequence of MxiH  
  
<220>  
<223> sequence can be found at MedLine accession number NP\_858270.1

<400> 7  
Met Ser Val Thr Val Pro Asn Asp Asp Trp Thr Leu Ser Ser Leu Ser  
1 5 10 15  
Glu Thr Phe Asp Asp Gly Thr Gln Thr Leu Gln Gly Glu Leu Thr Leu  
20 25 30  
Ala Leu Asp Lys Leu Ala Lys Asn Pro Ser Asn Pro Gln Leu Leu Ala  
35 40 45  
Glu Tyr Gln Ser Lys Leu Ser Glu Tyr Thr Leu Tyr Arg Asn Ala Gln  
50 55 60  
Ser Asn Thr Val Lys Val Ile Lys Asp Val Asp Ala Ala Ile Ile Gln  
65 70 75 80

Asn Phe Arg

<210> 8

<211> 252

<212> DNA

<213> *Shigella flexneri*

<220>

<223> nucleic acid encoding MxiH

<220>

<223> sequence can be found at MedLine accession number Nc\_004851

<400> 8

atgagtgtta cagtaccgaa tgatgattgg acattgagtt cattatctga aacttttgat 60

gatggaactc aaacattaca aggtgaacta acattggcac tagataaatt agctaaaaat 120

ccttcgaatc cacagttgct ggctgaatac caaagtaaatt tatctgaata tacattatat 180

aggaacgcgc aatccaatac agtgaaagtg attaaggatg ttgatgctgc aattattcaa 240

aacttcagat aa 252

<210> 9

<211> 73

<212> PRT

<213> *Escherichia coli*

<220>

<223> amino acid sequence of EscF

<220>

<223> sequence can be found at MedLine accession number NP\_312579.1

<400> 9

Met Asn Leu Ser Glu Ile Thr Gln Gln Met Gly Glu Val Gly Lys Thr  
1 5 10 15

Leu Ser Asp Ser Val Pro Glu Leu Leu Asn Ser Thr Asp Leu Val Asn  
20 25 30

Asp Pro Glu Lys Met Leu Glu Leu Gln Phe Ala Val Gln Gln Tyr Ser  
35 40 45

Ala Tyr Val Asn Val Glu Ser Gly Met Leu Lys Thr Ile Lys Asp Leu  
50 55 60

Val Ser Thr Ile Ser Asn Arg Ser Phe

<210> 10  
 <211> 222  
 <212> DNA  
 <213> Escherichia coli

<220>  
 <223> nucleic acid encoding EscF

<220>  
 <223> sequence can be found at MedLine accession number NC\_002695

<400> 10  
 atgaatttat ctgaaattac tcaacaaatg ggtgaagtag gtaaaacgct gagcgattct 60  
 gtgccagagt tacttaatag caccgatttg gttaatgacc ctgaaaaaat gttagagttg 120  
 cagtttgcgg ttcagcaata ttctgcttat gtaacgtag aaagtggaat gttgaaaacg 180  
 ataaaagatc tgggtctcaac catttctaac cgtagttttt aa 222

<210> 11  
 <211> 240  
 <212> DNA  
 <213> Yersinia pestis

<220>  
 <223> nucleic acid encoding HT-YscF; an example of a recombinant YscF

<220>  
 <221> misc\_feature  
 <222> (214)..(237)  
 <223> His-tag sequence

<400> 11  
 atgagtaact tctctggatt tacgaaagga accgatgatg caaacaagc ggttaatgac 60  
 tcgatagcag cattgaaaga taagcctgac aaccgcggcg tacttgctga cttacaacat 120  
 tcaattaata aatggtcggt aatttacaat ataaactcaa ccatagttcg tagcatgaaa 180  
 gacttaatgc aaggcatcct acagaagttc cactcgagc accaccacca ccaccactga 240

<210> 12  
 <211> 95  
 <212> PRT  
 <213> Yersinia pestis

<220>  
 <223> amino acid sequence of HT-YscF; an example of a recombinant YscF

<220>



<221> misc\_feature

<222> (88)..(95)

<223> His-tag

<400> 12

Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp Ile Ala Asp Leu  
1 5 10 15

Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp Asp Ala Asn Lys  
20 25 30

Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys Pro Asp Asn Pro  
35 40 45

Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys Trp Ser Val Ile  
50 55 60

Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys Asp Leu Met Gln  
65 70 75 80

Gly Ile Leu Gln Lys Phe Pro Leu Glu His His His His His  
85 90 95

<210> 13

<211> 5537

<212> DNA

<213> Artificial Sequence

<220>

<223> pMJ119

<220>

<221> misc\_feature

<222> (563)..(1375)

<223> kan encoded on complementary strand

<220>

<221> misc\_feature

<222> (3518)..(4597)

<223> LacI encoded on complementary strand

<220>

<221> promoter

<222> (4983)..(5000)

<223> T7 promoter

<220>

<221> misc\_feature

<222> (5001)..(5001)

<223> T7 transcription start

<220>  
<221> misc\_feature  
<222> (5072)..(5104)  
<223> T7 Tag

<220>  
<221> CDS  
<222> (5114)..(5401)  
<223> HT-YscF

<220>  
<221> misc\_feature  
<222> (5381)..(5398)  
<223> His-tag sequence

<220>  
<221> terminator  
<222> (5466)..(5512)  
<223> T7 terminator

<400> 13  
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60  
  
cagcgtgacc gctacacttg ccagcgcacct agcgcgccgt cctttcgtt tcttcacctc 120  
  
ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg 180  
  
gttccgattt agtgctttac ggcacctcga cccccaaaaa cttgattagg gtgatgggtc 240  
  
acgtagtggg ccacgcacct gatagacggg ttttcgccct ttgacgttgg agtccacgtt 300  
  
ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggctctattc 360  
  
ttttgattta taagggattt tgccgatttc ggctatttgg ttaaaaaatg agctgattta 420  
  
acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480  
  
tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540  
  
tccgctcatg aattaattct tagaaaaact catcgagcat caaatgaaac tgcaatttat 600  
  
tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat gaaggagaaa 660  
  
actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg attccgactc 720  
  
gtccaacatc aatacaacct attaatctcc cctcgtcaaa aataaggtta tcaagtgaga 780  
  
aatcaccatg agtgacgact gaatccgggtg agaatggcaa aagtttatgc atttctttcc 840  
  
agacttggtc aacaggccag ccattacgct cgtcatcaaa atcactcgca tcaaccaaac 900  
  
cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac 960  
  
aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat 1020  
  
tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg gggatcgcag 1080

|   |      |
|---|------|
| tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc ggaagaggca   | 1140 |
| taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgctac   | 1200 |
| ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat cgatagattg   | 1260 |
| tgcacactga ttgcccgaca ttatcgcgag cccatttata cccatataaa tcagcatcca   | 1320 |
| tgttggaatt taatcgcggc ctagagcaag acgtttcccg ttgaatatgg ctcataacac   | 1380 |
| cccttgattt actgtttatg taagcagaca gttttattgt tcatgaccaa aatcccttaa   | 1440 |
| cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga   | 1500 |
| gacccctttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg   | 1560 |
| gtggtttggt tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc   | 1620 |
| agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag   | 1680 |
| aactctgtag caccgcctac atacctcgct ctgctaatec tgttaccagt ggctgctgcc   | 1740 |
| agtggcgata agtcgtgtct taccggggtg gactcaagac gatagttacc ggataaggcg   | 1800 |
| cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttgagcgg aacgacctac   | 1860 |
| accgaactga gataacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga  | 1920 |
| aaggcggaca ggtatccggg aagcggcagg gtcggaacag gagagcgcac gagggagctt   | 1980 |
| ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag   | 2040 |
| cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg   | 2100 |
| gcctttttac ggttcctggc cttttgctgg ctttttgcgc acatgttctt tcctgcgtta   | 2160 |
| tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc   | 2220 |
| agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cctgatgcgg   | 2280 |
| tattttctcc ttacgcatct gtgcgggtatt tcacaccgca tatatgggtgc actctcagta | 2340 |
| caatctgctc tgatgccgca tagttaagcc agtatacact ccgctatcgc tacgtgactg   | 2400 |
| ggtcattggc gcgccccgac acccgccaac acccgctgac gcgcctgac gggcttgtct    | 2460 |
| gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag   | 2520 |
| gttttcaccg tcataccgca aacgcgcgag gcagctgcgg taaagctcat cagcgtggtc   | 2580 |
| gtgaagcgat tcacagatgt ctgcctgttc atccgcgtcc agctcgttga gtttctccag   | 2640 |
| aagcgtaaat gtctggcttc tgataaagcg ggccatgtta agggcggttt tttcctgttt   | 2700 |
| ggtcactgat gcctccgtgt aagggggatt tctgttcatg ggggtaatga taccgatgaa   | 2760 |

|             |            |             |            |            |            |      |
|-------------|------------|-------------|------------|------------|------------|------|
| acgagagagg  | atgctcacga | tacgggttac  | tgatgatgaa | catgcccggg | tactggaacg | 2820 |
| ttgtgagggg  | aaacaactgg | cggatatggat | gcggcgggac | cagagaaaaa | tcactcaggg | 2880 |
| tcaatgccag  | cgttcgtta  | atacagatgt  | aggtgttcca | cagggtagcc | agcagcatcc | 2940 |
| tgcgatgcag  | atccggaaca | taatggtgca  | gggcgtgac  | ttccgcgttt | ccagacttta | 3000 |
| cgaaacacgg  | aaaccgaaga | ccattcatgt  | tgttgctcag | gtcgcagacg | ttttgcagca | 3060 |
| gcagtcgctt  | cacgttcgct | cgcgtatcgg  | tgattcatto | tgctaaccag | taaggcaacc | 3120 |
| ccgccagcct  | agccgggtcc | tcaacgacag  | gagcacgac  | atgcgcaccc | gtggggccgc | 3180 |
| catgccggcg  | ataatggcct | gcttctcgcc  | gaaacgtttg | gtggcgggac | cagtgcagaa | 3240 |
| ggcttgagcg  | agggcggtga | agattccgaa  | taccgcaagc | gacaggccga | tcacgtcgc  | 3300 |
| gctccagcga  | aagcggtcct | cgcgaaaaat  | gacccagagc | gctgccggca | cctgtcctac | 3360 |
| gagttgcatg  | ataaagaaga | cagtcataag  | tgcggcgacg | atagtcatgc | cccgcgcca  | 3420 |
| ccggaaggag  | ctgactgggt | tgaaggctct  | caagggcatc | ggtcgagatc | ccggtgccta | 3480 |
| atgagtgagc  | taacttacat | taattgcgtt  | gcgtcactg  | cccgtttcc  | agtcgggaaa | 3540 |
| cctgtcgtgc  | cagctgcatt | aatgaatcgg  | ccaacgcgcg | gggagaggcg | gtttgcgtat | 3600 |
| tgggcgccag  | ggtgggtttt | cttttcacca  | gtgagacggg | caacagctga | ttgcccttca | 3660 |
| ccgcctggcc  | ctgagagagt | tgcagcaagc  | ggtccacgct | ggtttgcccc | agcaggcgaa | 3720 |
| aatcctgttt  | gatggtggtt | aacggcggga  | tataacatga | gctgtcttcg | gtatcgtcgt | 3780 |
| atcccactac  | cgagatatcc | gcaccaacgc  | gcagcccgga | ctcggtaatg | gcgcgcattg | 3840 |
| cgcgcagcgc  | catctgatcg | ttggcaacca  | gcacgcagc  | gggaacgatg | ccctcattca | 3900 |
| gcatttgcat  | ggtttgttga | aaaccggaca  | tggcactcca | gtcgccttcc | cgttccgcta | 3960 |
| tgggctgaat  | ttgattgcga | gtgagatatt  | tatgccagcc | agccagacgc | agacgcgccg | 4020 |
| agacagaact  | taatgggccc | gctaacagcg  | cgatttgctg | gtgacccaat | gcgaccagat | 4080 |
| gctccacgcc  | cagtcgcgta | ccgtcttcat  | gggagaaaat | aatactgttg | atgggtgtct | 4140 |
| ggtcagagac  | atcaagaaat | aacgcgggaa  | cattagtgc  | ggcagcttcc | acagcaatgg | 4200 |
| catcctgggc  | atccagcgga | tagttaatga  | tcagcccact | gacgcgttgc | gcgagaagat | 4260 |
| tgtgcaccgc  | cgttttacag | gcttcgacgc  | cgttcgttcc | taccatcgac | accaccacgc | 4320 |
| tggcaccacg  | ttgatcgggc | cgagatttaa  | tcgcccgcac | aatttgcgac | ggcgcgtgca | 4380 |
| gggcccagact | ggagggtggc | acgccaatca  | gcaacgactg | tttgcccgcc | agttgttgtg | 4440 |
| ccacgcggtt  | gggaatgtaa | ttcagctccg  | ccatcgccgc | ttccactttt | tcccgcgttt | 4500 |

|  |      |
|--|------|
| tcgcagaaac gtggctggcc tggttcacca cgcgggaaac ggtctgataa gagacaccgg  | 4560 |
| catactctgc gacatcgtat aacgttactg gtttcacatt caccaccctg aattgactct  | 4620 |
| cttcggggcg ctatcatgcc ataccgcgaa aggttttgcg ccattcgatg gtgtccggga  | 4680 |
| tctcgacgct etcccttatg cgactcctgc attaggaagc agcccagtag taggttgagg  | 4740 |
| ccgttgagca ccgccgccgc aaggaatggt gcatgcaagg agatggcgcc caacagtccc  | 4800 |
| ccggccacgg ggctgccac catacccacg ccgaaacaag cgctcatgag cccgaagtgg   | 4860 |
| cgagcccgat cttcccatc ggtgatgtcg gcgatatagg cgccagcaac cgcacctgtg   | 4920 |
| gcgccggtga tgccggccac gatgcgtccg gcgtagagga tcgagatctc gatcccgcgga | 4980 |
| aattaatacg actcactata ggggaattgt gagcggataa caattcccct ctagaaataa  | 5040 |
| ttttgtttaa ctttaagaag gagatataca tatggctagc atgactggtg gacagcaa    | 5100 |
| gggtcgggat ccg atg agt aac ttc tct gga ttt acg aaa gga acc gat     | 5149 |
| Met Ser Asn Phe Ser Gly Phe Thr Lys Gly Thr Asp                    |      |
| 1 5 10   |      |
| atc gca gac tta gat gcg gtg gct caa acg ctc aag aag cca gca gac    | 5197 |
| Ile Ala Asp Leu Asp Ala Val Ala Gln Thr Leu Lys Lys Pro Ala Asp    |      |
| 15 20 25   |      |
| gat gca aac aaa gcg gtt aat gac tcg ata gca gca ttg aaa gat aag    | 5245 |
| Asp Ala Asn Lys Ala Val Asn Asp Ser Ile Ala Ala Leu Lys Asp Lys    |      |
| 30 35 40   |      |
| cct gac aac ccg gcg cta ctt gct gac tta caa cat tca att aat aaa    | 5293 |
| Pro Asp Asn Pro Ala Leu Leu Ala Asp Leu Gln His Ser Ile Asn Lys    |      |
| 45 50 55 60  |      |
| tgg tcg gta att tac aat ata aac tca acc ata gtt cgt agc atg aaa    | 5341 |
| Trp Ser Val Ile Tyr Asn Ile Asn Ser Thr Ile Val Arg Ser Met Lys    |      |
| 65 70 75   |      |
| gac tta atg caa ggc atc cta cag aag ttc cca ctc gag cac cac cac    | 5389 |
| Asp Leu Met Gln Gly Ile Leu Gln Lys Phe Pro Leu Glu His His His    |      |